

디지털타임스 뉴스인쇄

기사 주소: http://www.dt.co.kr/contents.html?article_no=2010012802011860744001

[알아봅시다] 콘텐츠 필터링 시스템

불법복제 콘텐츠 전송 막아 유통 사전차단
웹하드 파일 업로드시 저작권침해 판별시스템 확인
검색 단어 제한·워터마크 삽입 등 다양한 기술 활용

박상훈 기자 nanugi@dt.co.kr | 입력: 2010-01-27 21:02



지난해 국내 인터넷TV(IPTV) 가입자 수는 150만명을 넘었습니다. 인터넷이 설치된 곳이라면 어디서든 최신 영화를 저렴하고 손쉽게 볼 수 있다는 점이 IPTV의 성공 요인 가운데 하나로 꼽힙니다.

그러나 IPTV의 확산은 동시에 콘텐츠 보호 측면에서 새로운 과제를 던져주고 있습니다. 과거에는 DVD 출시 이후 이를 파일로 만들어 유통하는 것이 주류를 이뤘다면 IPTV가 대중화되면서 콘텐츠를 복제해 배포할 수 있는 또다른 경로가 생긴 것입니다. 개봉관에서 상영을 마친 영화가 DVD 출시 이전에 IPTV를 통해 서비스되고 이것이 다시 파일로 만들어져 인터넷 파일공유 사이트나 웹하드 등에 등장해 논란이 일기도 했습니다.

이에 따라 최근에는 IPTV를 통한 콘텐츠 불법복제와 유통을 추적하기 위한 다양한 장치들이 연구, 적용되고 있습니다. 콘텐츠가 재생되는 시점에 디지털 워터마킹 기술을 이용해 콘텐츠의 사용자 정보를 삽입하고 불법 복제자를 찾아내는 포렌식 마킹(Forensic Marking) 기술이 대표적입니다.

그러나 포렌식 마킹 기술은 불법복제 후 유통되고 있는 콘텐츠에 대한 사후 대응적 방법이라는 점에서 콘텐츠 유통으로 인한 IPTV 사업자의 매출 감소를 막는 데는 직접적인 도움이 되지 못합니다. 대신 불법 복제된 콘텐츠가 유통되지 못하도록 하는 역할은 콘텐츠 필터링 기술이 맡고 있습니다.

콘텐츠 필터링이란 파일공유 사이트 등에서 콘텐츠의 저작권 침해 여부에 따라 파일전송을 제어하는 기술입니다. 크게 키워드 필터링, 해시(hash) 필터링, 특징점(feature) 필터링 등으로 구분됩니다.

먼저 키워드 필터링은 콘텐츠 파일명이나 제목 등 주요 단어와 텍스트를 이용하는 기술입니다. 영화 제

목, 드라마 제목, 배우 이름 등 사용자가 입력한 문자열이 저작물 제목과 같거나 흡사할 경우 검색이나 다운로드가 안되도록 막는 원리입니다. 키워드 필터링은 저작물 침해 여부를 빠르게 판단해 처리 속도의 강점이 있는 반면 파일명을 바꾸면 필터링 작업을 우회할 수 있습니다. 흔히 파일공유 사이트에서 음악콘텐츠 확장자를 mp3 대신 zip 등으로 바꾸면 공유할 수 있는 것도 이 때문입니다.

해시 기반 필터링은 파일마다 존재하는 고유값인 해시를 이용합니다. MD4, MD5, SHA 등이 대표적인 해시함수로 이들은 마치 인간의 지문처럼 파일마다의 고유한 코드가 됩니다. 해시 필터링은 불법 복제 콘텐츠의 해시값을 데이터베이스화해서 유통하는 각 단계마다 해시값과 비교해 다운로드, 업로드를 통제할 수 있습니다.

해시함수 역시 계산이 빨라 데이터베이스 구축이나 불복복제 여부가 빠른 장점이 있습니다. 그러나 콘텐츠 파일의 미세한 변화에도 해시값이 함께 변하기 때문에 압축포맷을 바꾸거나 파일 크기 등을 바꿔 일종의 변종파일을 만들어 유통할 경우 이를 지속적으로 데이터베이스에 업데이트해야 하는 단점이 있습니다.

특징점 필터링은 콘텐츠 고유의 특성을 추출해 불법복제 여부를 가려냅니다. 콘텐츠의 컬러정보, 명암정보, 모션정보, 주파수 대역 에너지값 등 다양한 정보를 이용하기 때문에 포맷이나 파일명을 변경해도 불법복제 콘텐츠 여부를 정확하게 파악할 수 있습니다. 반면 키워드 필터링이나 해시 필터링보다 계산해야 할 내용이 많고 특징점을 저장해야 하는 공간이 늘어나는 것이 단점으로 지적됩니다.

이러한 콘텐츠 필터링 기술을 이용한 필터링 시스템은 크게 저작권 침해 판별시스템, 웹하드 서비스 서버, 그리고 웹하드 서비스 사용자로 구성됩니다.

저작권 침해 판별시스템은 저작권 보호를 받고자 하는 콘텐츠를 이용해 콘텐츠 제목, 해시값, 특징점 등을 추출하는 역할을 하며 콘텐츠 대표정보 추출 프로그램, 추출 정보 데이터베이스, 웹하드 서비스 업체에 올라와 있는 콘텐츠 정보값과 저작권 데이터베이스를 검색해 침해 여부를 판별하는 프로그램 등으로 구성돼 있습니다. 이렇게 구축된 시스템은 웹하드 서비스 사용자가 불법공유 콘텐츠를 업로드하려고 하는 순간부터 작동됩니다. 즉 웹하드 서비스 업체의 업로드 서버에서 해당 콘텐츠의 정보값을 추출해 저작권침해 판별 시스템에 전송하면 여기서 침해 여부를 확인해 통보하고 이를 기준으로 콘텐츠 업로드 허용 여부를 실시간으로 보여줍니다.

현재 국내 대부분의 웹하드 서비스 업체들은 금치어를 설정해 검색을 제한하는 방식, 즉 키워드 필터링을 사용하고 있습니다. 각 기업들은 중복된 파일을 줄이기 위해 해시를 사용하지만 동일한 콘텐츠가 수십개의 불법 파일로 존재하기 때문에 충분한 필터링을 한다는 것이 현실적으로 불가능한 상황입니다. 가장 강력한 필터링 방식인 특징점 필터링은 데이터베이스 구축과 운영비용이 부담스러운 수준이어서 영세한 웹하드 서비스 업체들이 도입하기에는 쉽지 않은 상황입니다.

최근에는 특징점 필터링의 단점을 보완한 기술도 선보이고 있습니다. 워터마크를 이용한 콘텐츠 필터링 시스템이 그것으로, IPTV 콘텐츠에 저작권 관련 워터마크를 삽입하고 이를 검출할 수 있는 소프트웨어를 제작권 단체와 파일공유 사이트에 제공해 파일 업로드시 대조할 수 있도록 하는 것입니다. 이렇게 되면 특징점 필터링처럼 대규모 데이터베이스를 구축할 필요는 없지만 워터마크 검출 시간이 오래 걸릴 경우 웹하드 서비스 사용자들이 불편함을 느낄 수 있습니다.

전문가들은 콘텐츠 필터링이 앞으로 점점 더 강화될 것으로 전망합니다. IPTV 시장이 커지면서 개봉

관에서 상영을 마친 영화를 DVD 출시전에 IPTV에서 서비스하는 모델이 IPTV 업체의 새로운 수익원으로 떠오르고 있고 IPTV를 통한 불법 콘텐츠 유출은 자칫 기존의 DVD 시장마저 위협할 수 있기 때문입니다.

박상훈기자 nanugi@

자료협조=KT경제경영연구소

출력시간: 2011-01-18 10:45:01